ELSEVIER

Contents lists available at ScienceDirect

Science & Justice

journal homepage: www.elsevier.com/locate/scijus





A forensic international market survey of condom lubricants and personal hygiene products using ATR-FTIR coupled to chemometrics

Céline Burnier^{a,*}, Sally Coulson^b, Geneviève Massonnet^a, Kari Pitts^c, Georgina Sauzier^{d,e}, Simon W. Lewis ^{d,e}

- ^a Ecole des Sciences Criminelles, University of Lausanne, Switzerland
- ^b Institute of Environmental Science and Research Ltd, 120 Mt Albert Road, Sandringham, Auckland 1025, New Zealand
- ^c Forensic Science Laboratory, ChemCentre WA, Bentley, Australia
- ^d School of Molecular and Life Sciences, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia
- e Curtin Institute for Functional Molecules and Interfaces, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia

ARTICLE INFO

Keywords: Silicone Lubricants Glycerine International market Chemometrics ATR-FTIR

ABSTRACT

Condom residues may be encountered in forensic investigations as traces in sexual assault or rape cases. Casework studies have shown the value of distinguishing condom residues from other types of personal products used by women. However, up to now, there has been no investigation of their chemical variability within an international context. This work employed attenuated total reflectance Fourier transform infrared spectroscopy with chemometrics to provide objective characterisation of condom lubricants and personal hygiene products from the international market. 166 samples were obtained covering five major classes of products likely to be used by women. Principal component analysis distinguished most major classes based on their spectral profiles, with subsequent support vector machine models yielding discrimination accuracies over 90%. A two-step approach was subsequently developed and enabled both classification and a discrimination accuracy of 100%. This could provide greater confidence in chemical discrimination of residues from these products when conducting investigations and help assess the origin of the chemical profile obtained. Further testing using three validation sets produced an identification accuracy of 100% for generic classes, which may allow investigative leads to be more readily obtained from recovered evidence.

1. Introduction

Studies have reported the use of condoms in sexual assaults and rape cases, leading to forensic trace evidence potentially left in the vaginal matrix, on the penis of the perpetrator or in the victim's underwear [1-8]. The recovery and characterisation of these traces may provide evidence of association which can corroborate, or contradict, the allegations of the parties involved.

When condom evidence is investigated, either because the victim has indicated that a condom was used or does not remember the exact circumstances of the offence, the detection of lubricant trace evidence is of primary interest. Any resulting evidence will then be evaluated to determine if its source was indeed a condom. This question is of great importance in casework, as reported in the literature [2,4,9,10] and described hereafter.

In the United States, attempts have been made to answer this

question based mainly on analytical observations [2,5,10]. In these cases, positive (i.e. presence of silicone) [2,5,10] or negative (i.e. absence of silicone) [2] profiles were obtained, and in one case a comparison between a trace and a reference condom performed [5]. When the profile was positive, it was concluded that a condom was present, while a negative profile was taken to indicate the absence of the use of a condom [2]. Similarly, in the *Regina v. Andrew Nicholas Malkinson* case (2006) in the United Kingdom,[10,11] PDMS traces were detected on the victim's underwear. The victim's statement was that she never used personal products. Therefore, the court stated that if the victim's statement was correct, it was more likely that the traces originated from a condom [10,11].

However, indistinguishable profiles between the trace and the source does not mean that the reference sample is the source of the trace evidence. At the time of the publication of these cases, i.e. between 1994 and 2006, there was no published model that identified the error rates

^{*} Corresponding author at: Ecole des Sciences Criminelles, Quartier UNIL-Sorge, Bâtiment Batochime, CH-1015 Lausanne, Switzerland. E-mail address: celine.burnier@unil.ch (C. Burnier).

related to these analyses, nor that allowed interpretation of a positive nor a negative result. False positives generated by hygiene products or alterations due to textiles have not been investigated, and these cases could have benefited from more research regarding the discrimination of condom chemical profiles using similar instrumental analysis, i.e. Fourier transform infrared spectroscopy.

The aforementioned casework in addition to earlier studies reported the use of vibrational spectroscopy as a non-destructive means to characterize a trace. Several such studies have used infrared spectroscopy as a screening method to examine condom lubricants [2,9,12-15] and classify them as a function of their viscosity [13], while Raman spectroscopy has also been used for trace evidence detection [3,16]. As these methods are recommended for identifying potential condom residues in casework [2,9], it is of prime importance to define the discriminating capabilities of the technique to face issues of sample differentiation or source identification. One common way to evaluate discriminatory performance is to conduct a statistical exploratory study on a large population set. Although a recent study found that DRIFTS-FTIR [15] showed greater potential for characterising lubricants, ATR-FTIR was chosen for this study as it readily available in most forensic laboratories and requires minimal sample preparation.

This preliminary study made use of bulk samples (i.e. personal hygiene products, condom lubricants and intimate lubricants) found on the Australian, New Zealand and Swiss markets, in order to first evaluate any structure in the market and assess the ability of ATR-FTIR to distinguish bulk lubricants based on their chemical profile. The spectra obtained were qualitatively analysed based on visual observation before being subjected to chemometrics. As highlighted in both the National Academy of Sciences (NAS) [17,18] and more recently in the PCAST [19] report, visual interpretation of forensic evidence can suffer from operator bias or subjectivity. The use of chemometrics can minimise this bias as it uses mathematical and statistical methods to explore differences in a given dataset. Chemometrics is also known to be useful for pattern recognition and may help to reveal trends not clearly visible otherwise. This is useful in observing the structure in the data and understanding the structure of the market.

Assertions on the content of the samples present on the market have never been verified and the discrimination of similar chemical profiles have been verified only on a small set of samples. Previous papers [9,12,20-22] offered an overview of data acquired on a small set of samples, rarely going over 100 samples. The present paper offers conclusions drawn on one of the biggest known datasets regarding condom and personal hygiene products. The significant amount of data and their variable provenance allows more accurate conclusions to be drawn, as a benefit of a more thorough representation of the market. This paper offers an international perspective, which may help the understanding of the worldwide market.

2. Material and methods

2.1. Samples

100 condoms representing 8 brands present on the New Zealand market were purchased from major distributors and manufacturers. 20 condoms representing 6 brands present on the Swiss market were purchased from Swiss supermarkets and pharmacies; and 6 condoms and representing 4 brands present on the Australian market were purchased from Western Australian supermarkets and pharmacies, for a total of 127 condoms. 6 creams, 5 personal hygiene products, 1 massage oil and 26 lubricants that may be used by women on a daily basis, representing 9 brands from the Australian and international market, were purchased from pharmacies. The samples obtained were considered representative of the market share of the major condom and personal hygiene brands and sub-brands available to consumers. In total, 166 samples were analysed. The entire listing of the samples is available in Appendix A.

Reference material was used for the identification of the main

expected compounds: PDMS 200cSt and and Polyethylene Glycol (PEG) 400 were purchased from Sigma-Aldrich, Glycerin from Univar, Ajax Chemicals and used as received.

2.2. Instrumental conditions and data acquisition

Infrared spectra were collected using a Nicolet iS50 FTIR spectrometer equipped with single-bounce diamond crystal ATR accessory (Australian samples), Smart Orbit attachment (New Zealand samples) or Golden Gate Single Reflection Diamond ATR system (Swiss samples). Data collection was carried out using the OMNIC software. Spectra were collected over the 4000 to 400 cm⁻¹ range with 4 cm⁻¹ resolution and 32 co-added scans. ATR correction was performed on all spectra to account for variations in penetration depth based upon wavelength.

Condoms were rubbed directly on the ATR crystal and analysed with no further preparation. All other products were applied as thin films to cover the ATR crystal and analysed with no further preparation. The sampling window was thoroughly cleaned using ethanol and lint-free tissue before each sample, and a background scan of the clean crystal obtained between each replicate acquisition. For each sample, 5 replicates were acquired, to be able to statistically consider the sample variation. In a single sample, if there was variation amongst the 5 replicates and more than one chemical profile could be observed, an additional 5 replicates were run to ensure adequate representation of this variability. A total of 830 analyses were used for the present study.

2.3. Data analysis

2.3.1. Qualitative analysis

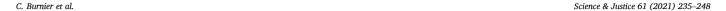
Qualitative observations were carried out using OMNIC software and samples were visually classified into different groups according to their chemical profiles. Data pre-processing and chemometric analysis were performed using the Unscrambler® X 10.5 software (Camo Software AS, Oslo, Norway). Spectra were truncated to omit the 2340–1880 cm⁻¹ region due to interference from the diamond crystal. Then range normalisation was applied to remove variation related to the amount of sample deposited and therefore the sample layer thickness on the crystal. Before performing the complete exploratory study, it was important to know whether consistent results were obtained when analysing the same types of condoms. Consistency was evaluated using the replicates and visually assessing that peak position, peak shape and abundance were similar within a same condom (same brand, model, lot number). Tests, using discriminant analysis, were performed to evaluate if different lots of the same brand presented different chemical profiles.

2.3.2. Chemometrics

Using the Unscrambler X (v. 10.5), principal component analysis (PCA) was carried out using mean-centered spectra and non-linear iterative partial least square (NIPALS) algorithm, with 1000 iterations. Samples were plotted using up to the first three principal components (PCs) to visualise the distribution within the samples and identify any clustering. 12 different pre-processing methods (available in Appendix B) were applied and the resulting PCA plots compared to see which combination allowed the best visual discrimination of the samples.

2.3.3. Discrimination model creation and comparison

Given that visual observation of principal component clustering can still suffer from operator's subjectivity, supervised classification was used to provide more objective discrimination. Models were constructed using the first three PCs, treating each replicate of each sample as a separate sample, as they were analysed as if they came from different samples. Discrete classes were attributed to each replicate based on the observations of the chemical profile and the knowledge of the samples (i. e. sample type, sample content, brand, model). Five classification algorithms; linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and support vector machine (SVM) with linear,



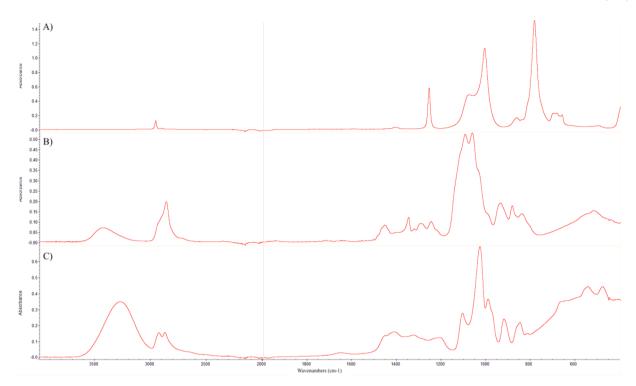


Fig. 1. Reference FTIR chemical profiles used to visually discriminate spectra from the overall dataset, a) silicone-based, b) PEG-based c) Glycerine-based.

polynomial (3rd degree) and radial basis function algorithms, were used and compared. Confusion matrices showing the actual and predicted classes were obtained, and model performances were calculated for each discrimination model created. The predicted classes using the test set were compared to the actual class, to evaluate the accuracy of the model using performance analysis parameters, such as classification error rate, false positive and false negative rate.

2.3.4. Discrimination model validation

Performance on an external sample set is crucial to rigorous validation, as using the same data to build and test a model can result in overly optimistic discrimination and classification accuracies. Therefore, the different models were evaluated using an external sample set and the performances were calculated. The models were then tested on three external sample batches involving different analysts and different instruments: a known-matching samples set (i.e. a batch of samples already represented in the model); a known non-matching samples set (i. e. a batch of samples not yet represented in the model); and a blind validation consisting of samples unknown to the analyst. Regarding the blind validation procedure, and a second operator acquired new data, from selected samples taken in the dataset, and the results were interpreted by another operator. Such an approach reinforces the quality of the model as experimental factors may significantly affect the observation, due to instrumental and operator variation. Therefore, the higher the correct classification, the stronger the model is for further applications. Samples were subjected to the same data pre-processing as the samples used to build the model and were then projected and classified using each model.

3. Results and discussion

3.1. Preliminary considerations

Suitable replicates are essential in a forensic context. In the case of this study, replicates not only allow consideration of sample or instrumentation variability, but also operator variation such as in depositing samples on the ATR crystal. It was observed during the analysis that,

when analysing non-condom samples, if the droplet on the crystal was too thick then minimal spectral variation was seen between samples. Water-based chemical profiles would be observed, and the profile would change with water evaporating. Therefore, the deposition strategy to ensure proper qualitative observation of a full and adequate chemical profile was to deposit the sample on a gloved finger and then rub it on the crystal to ensure that the sample was applied as a thin film.

As previously reported in the literature, [9,12,23] at least 3 major classes, i.e. silicone-based, water-based, or oil-based (vegetable or mineral oils), are most likely to be encountered on the market. Within the water-based class, the most common lubricants are PEG and glycerine. Reference materials were run to observe the differences between the different lubricants. Chemical profiles from PDMS, PEG and glycerine are presented in Fig. 1. PEG and glycerine were distinguished based on the peak number and position, as well as the OH peak shape. For the oil content, comparison with ATR spectra found in the literature and NIST webbook online database (https://webbook.nist.gov/chemistry) was used.

Any sub-groupings within these main categories could be due either to variation in the deposition and analysis approach or to variation in the sample composition. The use of replicates, in this case, is needed to ensure that any observed characteristics are consistent in all the profiles and hence differences are attributable to compositional changes rather than variation in the analysis.

These experiments aimed to study if samples coming from condoms and other sources could be distinguished, as well as to evaluate the potential discrimination of different types of condoms.

Chemical profiles of 10 Durex Extra Safe condoms were not found to present distinguishable chemical profiles, and the observation of the PCA scores plot (Appendix C) did not highlight differences within the samples presenting different lot number. These observations give confidence that there is no observable difference between production lots.

3.2. Chemical qualitative classification

Qualitative spectral identification was carried out to determine whether different chemical profiles were observed in the overall data

Table 1Categories of chemical profiles obtained from qualitative analysis of the dataset, along with the number of samples in each group or subgroup, their characteristic infrared absorption frequencies and corresponding components identified through comparison with reference materials.

Group	Number of samples	Sample no	Infrared Absorption Frequencies (cm- ¹)	Components
1	129		790, 1012, 1090, 1260, 2960	PDMS
1a	124	1-13; 15-18; 20-21; 23-56; 58-69; 71-100; 109-111; 113-123; 125-133, 169,171, 178-181	790, 1012, 1090, 1260, 2960	
1b	3	14, 19, 22	790, 865, 1012, 1090, 1260, 2960	
1c	1	174	700, 790, 1050, 1260, 2960	
1d	1	156	790, 1012, 1090, 1260, 2960	
2	15		850, 921, 992, 1029, 1108, 1410, 2879, 2936, 3280	Glycerine
2a	9	57, 70, 108, 175, 182, 165, 172,176, 177	799, 839, 922, 992, 1036, 1108, 1261, 2879, 2936, 3310	
2b	5	135, 137, 140, 141, 142	922, 992, 1040, 1112, 1414,1641, 2888, 2944, 3320	
2c	1	163	994, 1043, 1112, 1639, 2848, 2916, 3345	
3	2	138, 158, 159, 173	722, 1098, 1160, 1237, 1464, 1743, 2853, 2922, 3008, 3408	Oils
4	1	124	519, 840, 885, 935, 1063, 1095, 1248, 1294, 1348, 1459, 2865, 3430	PEG
5	3	112, 168, 170	802, 837, 921, 990, 1040, 1078, 1136, 1640, 2881, 2935, 2975, 3330	Unknown water-based component
6	13	160, 161, 162, 164, 157, 154, 134, 136, 166, 167, 155, 132, 139	1636, 2853,2924, 3350	Water- dominated spectra

and if it was possible to visually classify them. Spectra were overlaid in Omnic software, and evaluation was based on the presence of peaks as well as their number, position and shape. Systematic comparison with reference materials was performed to ensure correct identification of the composition.

Five main categories representing different chemical profiles were obtained based on the chemical composition. The first four were silicone-based, oil-based, glycerine-based, and PEG-based as expected. Category 5 is made of an unknown water-based component presenting a different chemical profile than glycerine and PEG. An additional sixth category gave a chemical profile of products dominated by water, which were hence unable to be reliably assigned to any of the specific water-based sub-groups. Spectra were then examined within each category, and sub-groups within the 2 major groups (silicone-based and glycerine-based), were identified (Table 1). Subgroups were defined based on peak presence or absence, peak position and peak shape. Table 1 lists the number of samples in each category, and the IR bands characteristic of each chemical profile.

Silicone-based samples (Group 1 in Table 1, Fig. 1a) were the main

category observed with the presence of characteristic peaks associated with PDMS corresponding to symmetric and asymmetric Si-O stretching at 1020 and 1090 cm^{-1} , a Si-C stretching at 1263 cm^{-1} and a C-H dimethyl and trimethyl deformation around $800~{\rm cm}^{-1}$ [15]. Group 1 samples only presented these four peaks, thus suggesting that either PDMS is not mixed with any other component when applied to a condom or in a lubricant, or PDMS is predominant in the spectrum and other components are present in too low a relative amount to be detected. Subclusters 1a to 1d were created based on the observation of variations within the chemical profiles (see Appendix D). Category 1a was the most commonly observed profile and was matching the reference spectrum of PDMS, without any noticeable differences. Other sub-clusters were distinguished by the presence of additional peaks to the classic silicone pattern (1b), a single peak instead of a doublet for the silicone double bond (1c), or an inverted relationship between symmetric and asymmetric Si-O vibrations (1d). Differences observed in group 1b may be explained by the presence of other components:GC-MS analysis of the samples constituting group 1b revealed the presence of benzocaine and traces of PEG [24]. Regarding the differences within group 1c and 1d, a different viscosity of the PDMS are likely to generate such variations. This was confirmed by leading a pyrolysis-GC/MS analysis [data not shown1.

Oil-based samples (Group 3) did not present any typical vibration of O–H or silicone bonds, whereas all other sample types presented a strong O–H peak between 3700 and 3000 cm⁻¹. This peak was found in all samples containing PEG (Group 4), glycerine (Group 2) or more generally water-based products (Groups 5–6). However, differences between the O–H peak shape as well as additional peaks allowed these samples to be separated into different categories (Appendix D). The presence of additional peaks as indicated in Table 1, for example peaks at 800, 1261, or 1640 cm⁻¹ in the glycerine group, allowed the creation of sub-categories within this group. These peaks are potentially due to other water-based additives.

A more careful observation of the different groups highlighted that water-based type samples (Groups 2, 4, 5 and 6) did not present any strong vibrations linked to silicone-based products. Likewise, silicone-based samples did not present any peaks specific from glycerine- or PEG-based lubricants. These observations suggest that condom lubricants appear to be either silicone-based or water-based, but not a mixture of the two. This is surprising considering that some of the silicone-lubricated condoms are flavoured or known to contain additives dedicated to long-lasting pleasure or that some samples, such as Ceylor Gold samples contain both silicone and water-based compounds [24]. In addition, Maynard *et al* in 2001 [9] reported that some silicone-lubricated condoms or lubricants were also containing nonoxynol or glycerine. Such differences might be explained by the difference of the spectroscopic instrumentation used for the study as well as a modification of the market within the last 20 years.

3.3. Statistical discrimination of the samples

12 pre-processing methods were considered in this study, based on previous studies reporting the importance of pre-treatment on the discrimination of samples [25-29]. PCA was carried out for each set of pre-processing, but none were found to improve sample discrimination compared to the initial corrected dataset (baseline correction and range normalisation).

PCA was performed on the entire dataset without further preprocessing to visualise any clusters. The scores plot obtained using the first three PCs (accounting for 93% of the variation) is shown in Fig. 2. Silicone-based samples were separated from the water-based samples along PC1 and the oil-based samples along PC2. This clustering pattern reinforces the main categories identified during the qualitative examination of the spectra. Water-based samples were found to offer a strong variation along PC1 and PC2, revealing at least 2 sub-clusters. The use of PC3 showed that water-based samples were clustered together, but PEG

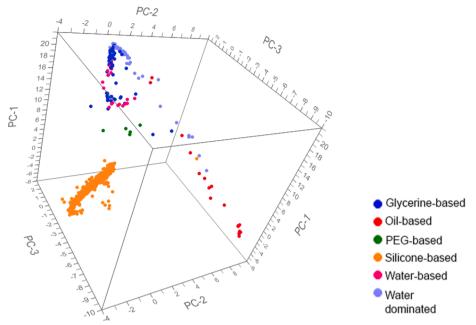


Fig. 2. 3-dimensional PCA score plot showing the distribution of the dataset according to the detected content obtained with ATR-FTIR spectra. Silicone-based samples (Group 1) are represented by the orange dots, glycerine-based (Group 2) samples in blue, Oil-based (Group 3) in red, PEG-based (Group 4) in green, water-based (Group 5) in pink and Water-dominated spectra (group 6) are in light violet. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

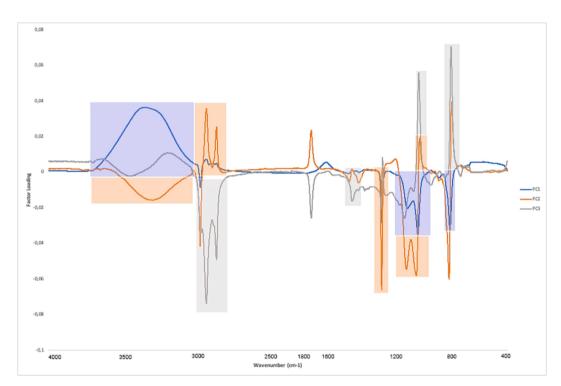


Fig. 3. Factor loadings of PCs 1-3 for PCA conducted on the entire dataset, based on their ATR-FTIR spectra.

containing samples were separated from the rest of the water-based samples.

Interestingly, visually distinguished samples were clustered in the same groups along these PC1 and PC2, and samples that were visually indistinguishable based on their spectra were found to be visually clustered in different groups in the scores plot. Additionally, different sub-groups could be observed compared to the visual examination. The minor peaks enabling visual differentiation between specific groupings were likely dismissed by PCA as having low significance and thus incorporated into later PCs that weren't examined. These observations demonstrate the importance of visually examining the data, and not just

 $relying \ on \ chemometric \ interpretation.$

The factor loadings for these PCs (Fig. 3) were used to identify the spectral regions, and thus specific chemical components, contributing to sample discrimination. PC1 was found to be negatively correlated with peaks at \sim 790, 1020, 1090, 1260 and 2963 cm $^{-1}$, which are characteristic of the silicone backbone from PDMS [9,12]. A highly positively correlated peak at \sim 3000–3700 cm $^{-1}$ was also observed, consistent with the O–H stretching bond attributed to water and water-based components [9,12]. Consequently, the discrimination between classes across PC1 is due to the presence or absence of silicones and water containing components. The samples containing silicone bonds attain

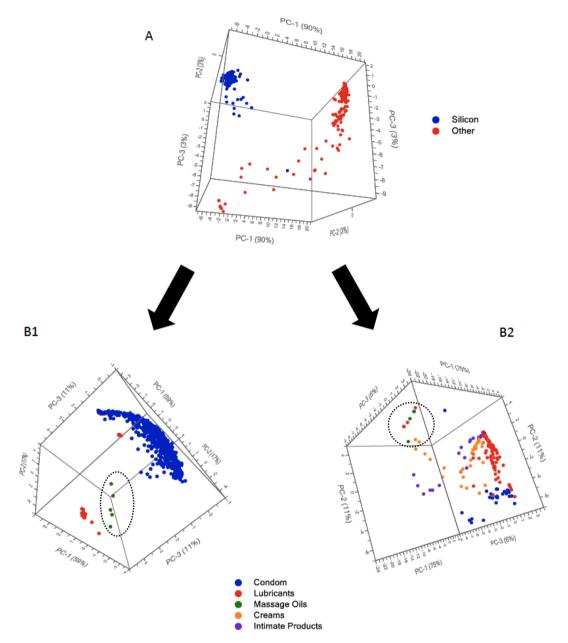


Fig. 4. 3-dimensional PCA score plot showing the distribution of the samples A) according to the chemical profile (i.e. silicone vs other), B1) within the silicone-containing cluster, B2) within the water-based containing cluster. Dotted circle highlight massage oils clusters in both populations.

large negative values on PC1 while the remaining samples attain significant positive score on PC1. These observations match the clustering that was observed in Fig. 2.

The same correlations appeared along PC2 (Fig. 3), suggesting that variations were detected within the silicone content of the different samples. Several other peaks were found to contribute to separation along PC2. A peak at 1000 cm⁻¹ was linked to glycerine, while those at 3700–3000 cm⁻¹ (O–H stretching), 2920 cm⁻¹ (CH₂ asymmetric stretching), and 2850 cm⁻¹ (CH₂ symmetric stretching) could be attributed to PEG, glycerine and other water-based components. This explains why water-based containing samples are all within the same interval along PC2 in Fig. 2. The peak at 1740 cm⁻¹ (C = O stretching) could not be attributed to any of the major compounds or other water-based components, suggesting it could be coming from an oily compound in some of the oil-based samples, as this vibration is usually associated with a carboxylic acid. This explains the strong variability of the water-based and oily samples. Finally, along PC3, peaks at 1254 cm⁻¹ (CH₂ twisting) and around 1360 and 1460 cm⁻¹ attributed to CH₂

wagging and CH₂ scissoring, allowed the discrimination of the PEG and glycerine containing samples on Fig. 2.

PCA scores plots showed that excessive leverage on the overall model was obtained due to silicone or non-silicone content (Fig. 4A), potentially reducing discrimination between the rest of the samples. Performing PCA separately on each of the two main categories was found to give additional information on the separation of the samples.

The silicone-based cluster was found to incorporate three classes of samples, including condoms, lubricants and massage oils. This observation highlights that silicones can be used in various other products that might be found in the vaginal matrix. Despite the improved separation of the silicone-containing samples, some overlap was observed between lubricants and condoms with a non-differentiated silicone chemical profile (Fig. 4B1). The present observations confirm previous assessments on the separation of condom lubricant types [9,13,23], as most of the condoms used in this study were found to share the same qualitative profile, i.e. silicone-based. Although lubricants were found to mainly contain water-based compounds, 18% (5 of 27 samples) of the

lubricant batch (Astroglide Waterproof Silicone Liquid, Ansell Skyn Maximum performance, Ansell Lifestyles Luxe Silicone Based Lubricant, Astroglide Diamond Silicone Gel Personal Lubricant, Durex Play Perfect Glide) presented a silicone-based profile that was clustered with the condom profiles. This was expected as these products specifically indicated the presence of silicone in their composition.

In the context of *R. v. Andrew Nicholas Malkinson* [8,10], this suggests that a silicone-based profile could originate from products other than condoms. It was also observed that 3.9% of the whole condom batch (5 of 126 samples) presented a significantly different chemical profile that clustered with water-based samples (Fig. 4B2). These condoms (*Ansell-Lifestyles Party Variety Warm/Cool, Ansell-Lifestyles Party Mix Warm Smooth, Manix OrgazMax Plus, Manix Endurance* and Ceylor Gold) represented 'warming' condoms from a single supplier (Manix being manufactured by Ansell) and potentially contained nonoxynol-9 (a spermicide) or other water-based additives, explaining their similarity to other water-based products. The other lubricants clustering with those special condom types were not warming lubricants.

Massage oils were found to present two different chemical profiles (Fig. 4B1 and Fig. 4B2), clearly separated from the other samples. It was interesting to note that a cluster was observed with other water-based lubricants (dotted circle on Fig. 4B1, and 4B2). It appeared that the lubricant in question was named as a Personal Lubricant and Massage Oil. It can be therefore assessed that massage oils, named as such, present a separate and specific profile, which can be separated from other types of samples.

Four of the six cream samples reported silicone in their composition but were not found to cluster close to silicone-based condoms. This may be explained by the complete composition being a mixture of glycerine and silicone-based components, with a predominance of glycerine, water, wax or paraffin wax leading to a specific profile. The two samples reporting no silicone in their composition were found to cluster very closely to the main lubricants class. This observation was expected, as the discrimination of the samples generated in the PCA was mainly dictated by the silicone composition.

Personal hygiene products, classified in Fig. 4 as intimate products, were generally grouped and close to cream and lubricant products. An exception was observed for one of the samples (FemFresh Feminine deodorant spray) which was found to be silicone-based. The initial composition of the product stated dimethicone as the major component of the product, dimethicone being another denomination for PDMS. Such differentiation was therefore not surprising. The dispersion on the water-based side of the clusters was due to only one sample (FemFresh Daily Intimate Wash,), whose profile was found to present a high variation within the replicates. Such variation can be explained by the variation of water evaporation after the deposition of a thin layer of a water-based product.

3.4. Structure of the international market

Studying the structure of the market aims to see if the clusters observed can be associated with particular characteristics of the samples. This is useful to generate investigative leads from a questioned sample by predicting what its origin might be. The main interest was to understand if it was possible to distinguish samples from different brands or models based on their infrared chemical profile.

The analysis revealed that no major differences were observed between the different markets in terms of the country of purchase. This was not surprising as the country of purchase does not necessarily correspond to a specific manufacturer.

To highlight potential variation between different brands, a smaller dataset containing only silicone-based condoms was used, as they were the biggest sample group in this study. Separate analysis did not highlight any evident clusters or separation of silicone-based condoms from the different brands. This suggests that different brands use the same type of chemicals for any given type of samples. These assumptions are

in agreement with earlier research [13], which concluded that most lubricated condoms present in the market used the same type of lubricant, with chemical properties being too similar to allow visual or statistical differentiation.

Within-brand analysis and lot number production variation were investigated using samples from *Durex*, the most common brand found on the international market [9,12,13,22,30,31]. No evident clusters were formed from analysing the spectra from different condom models within this brand, nor were any clusters highlighted regarding different lot numbers from the same brand and same model condom type (Appendix D). It thus appears that individual brands use the same lubricant composition across a range of products, and that there do not appear to be substantial variations in this composition across production lots.

3.5. Discrimination model and classification procedure

Five different discrimination models (LDA, QDA, and SVM with three different algorithms: linear, 3rd degree polynomic and radial basis function) were constructed based on the sample class. SVM analysis with a radial basis function kernel (RBF), γ (kernel parameter) set at 0.01 and C (soft margin parameter) set at 0.001 was found to present the highest discrimination (calibration) accuracy (92.3%) and predictive (validation) accuracy (92.05%), as well as the best classification regarding sample content (Appendix E). The model was constructed using the training set containing 2/3 of the dataset, and then the validation set (1/3 of the dataset) was predicted to provide a more realistic estimate of model performance. Only 3 samples out of 101 were misclassified; 2 false negatives (6% rate) and 1 false positive (13.3% rate).

Regarding false positives and false negatives, a false positive indicates that a profile from a personal hygiene product has been associated with a condom profile. In contrast, a false negative indicates that a profile from a condom has been associated with a profile from another intimate product. The interpretation of these two results in a real case is important. Observations of the classification quality showed that samples are easily classified according to their chemical profile, but that the identification of the membership of a sample class is difficult, has a high error rate and is certainly related to the initial classification by chemical profile. This suggests that classification should be undertaken in a two-step approach to limit errors:

- 1. Identify the sample caegory according to the chemical profile of silicone or non-silicone type (Fig. 4A). This step aims to determine whether silicones are present in the sample.
- Identify the class of origin separately within each population, i.e. within the silicone profile population or the other profile population (Fig. 4B1, Fig. 4B2).

Discriminant models were built using a two-step approach to classify the samples, with SVM again found to give the most accurate classification. This discriminant model gave 100% accuracies for both the calibration and validation sets. The model was then used to predict external test sets, including both samples present in the model and samples not yet represented in the model, analysed using different instruments and analysts. Finally, a blind validation test was carried out to validate and evaluate the robustness of the developed approach. From a forensic point of view, blind validations are mandatory before considering any application to casework, as it is representative of the simulation of a real scenario, where the source and the nature of the evidence are unknown. Also, it can minimise potential confirmation bias from a classic validation procedure.

In the present study, 100% correct classification of all external test samples was obtained using the two-step model. It was possible to classify samples according to their chemical content and their class. This is of great value for further application to casework as shown in [32], more specifically for investigative purposes, as it means that it is possible to infer the source of recovered forensic evidence. However, in

respect to casework application, further work is required to validate these findings using trace evidence samples, as characterisation may be affected by the presence of a biological matrix or other factors as discussed in previous studies [23,33].

4. Conclusion

This study investigated the distribution of condoms, lubricants, massage oils, creams and personal hygiene products from an international perspective through the analysis of 166 samples purchased between 2018 and 2019 from three different international markets.

At least 6 different chemical profiles could be visually and qualitatively differentiated. However, different clustering patterns were observed using unsupervised statistical analysis, highlighting this as a complementary method to qualitative examination. It is thus important to still visually examine the chemical profile of the sample rather than relying solely on statistical analysis. Condoms, lubricants, creams and other intimate products could be distinguished based on their qualitative profiles, although some condoms and lubricants were found to be atypical for their category. Nevertheless, the results indicate that condoms and other personal products can present distinct chemical profiles. As expected, no differentiation was observed according to the international market, brand, specific product or production lot.

ATR-FTIR spectroscopy with SVM was found to be a valuable tool for the discrimination and classification of samples according to their chemical profiles. The presented model allowed discrimination of samples according to whether they contained silicone or not, and whether it was a condom or another type of sample. It was observed that over 96% of the condom samples contained silicones as lubricants, but silicones were also detected in some personal hygiene products. As stated in the introduction, a key forensic question in cases of rape or sexual assault is whether an observed profile comes from a condom or another source. The observations in the present paper highlighted that it is possible that a silicone-containing trace did not originate from a condom. These findings have significance in the context of cases such as the *R. v. Andrew Nicholas Malkinson* case (2006) [8,10], where the ability to distinguish condom lubricants from other personal products could affect the interpretation of evidence within the context of the case. Interpretation of the

absence of evidence should be handled with great care, and the present paper does not allow a position on this precise point. False positives due to silicone-containing personal products may appear in casework samples. Victim interviews are necessary to interpret the evidence and complementary analysis, such as py-GC-MS, might be necessary to confirm the results of the FTIR analysis.

Initial discrimination and classification based on a single model presented a high level of false positive errors (up to 15%). Therefore, a two-step approach is recommended for the classification of infrared spectra; the first step is to identify if the sample contains a silicone-based profile or not, and the second step is to further classify the sample in the suitable model based on the initial profile. This methodology was successfully applied to predict the classes of known match samples, known non-match samples and blind validation samples representing unknown casework samples. ATR-FTIR spectroscopy is hence a powerful screening method for the investigation of condom or lubricant in forensic casework. Further research should investigate the classification of real traces, to see whether they classify the same way pure lubricant samples do

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge Curtin Institute for Functional Molecules and Interfaces for providing funds for sample collection, Peter Chapman (Curtin University) for assistance with the instrumentation, and Virginie Favre (Université de Lausanne) for assistance in data acquisition.

Fundings.

This research was partially funded by Fondation pour l'Université de Lausanne, who covered travelling grants. Funding towards purchase of Australian sourced samples was provided by the Curtin Institute for Functional Molecules and Interfaces.

Appendix A:. List of the samples used in the study

Abbr. Sign: NZ = New Zealand, AUS = Australia, CH = Switzerland

Sample	Brand	Product	Lot	Purchased
1	Durex	Extra Safe	1,000,026,318	NZ
2	Durex	Extra Safe	1,000,010,957	NZ
3	Durex	Extra Safe	1,000,106,077	NZ
4	Durex	Extra Safe	1,000,049,207	NZ
5	Durex	Extra Safe	1,000,046,754	NZ
6	Durex	Extra Safe	1,000,041,595	NZ
7	Durex	Extra Safe	1,000,049,636	NZ
8	Durex	Extra Safe	1,000,041,595	NZ
9	Durex	Extra Safe	21,306,074	NZ
10	Durex	Classic	1,000,063,396	NZ
11	Durex	Unknown	1,000,010,958	NZ
12	Durex	Pleasure Me	10,939,364	NZ
13	Shield	XL	PN29803	NZ
14	Durex	Mutual Climax	1,000,026,346	NZ
15	Durex	Unknown	1,000,015,117	NZ
16	Durex	Unknown	1,000,011,869	NZ
17	Durex	Unknown	1,000,012,759	NZ
18	Durex	Classic	10,121,993	NZ
19	Durex	Mutual Climax	1,000,039,272	NZ
20	Durex	Intimate Feel	1,000,044,392	NZ
21	Durex	Intimate Feel	10,837,252	NZ

(continued on next page)

(continued)

Sample	Brand	Product	Lot	Purchased
22	Durex	Mutual Climax	1,000,039,272	NZ
23	Durex	Confidence	1,000,059,628	NZ
24	Durex	Banana	14F2456B	NZ
25	Durex	Pleasure Me	1,000,039,189	NZ
26	Durex	Pleasure Me	1,000,042,518	NZ
27	Durex	Confidence	1,000,045,624	NZ
28	Durex	Unknown	22B09482	NZ
29	Durex	Unknown	1,000,049,211	NZ
30		Unknown		NZ NZ
	Durex		1,000,064,796	
31	Durex	Classic	10,135,039	NZ
32	Durex	Strawberry	14F1293S	NZ
33	Durex	Apple	14F1293A	NZ
34	Ansell	Contempo-Rough Rider	1,011,081,616	NZ
35	Durex	Strawberry	14F2456S	NZ
36	Durex	Orange	14F2456O	NZ
37	Durex	Classic	21,306,193	NZ
38	Durex	Thin Feel	10,847,195	NZ
39	Durex	Orange	14F1293O	NZ
40	Durex	Confidence	1,000,179,693	NZ
41	Gold Knight	Chocolate	PC29801	NZ
	· ·			
42	Marquis	Flavoured	PG1202	NZ
43	Durex	Extra Safe	1,000,136,970	NZ
44	Marquis	Regular	1,005,096	NZ
45	Gold Knight	Strawberry	PS2980	NZ
46	Durex	Extra Safe	1,000,174,930	NZ
47	Ansell	SKYN-Original	1,705,403,516	NZ
48	Ansell	Lifestyles- Ultra Thin	1,504,111,016	NZ
49	Ansell	Lifestyles- Ultra Thin	1,611,060,216	NZ
50	Ansell	Lifestyles- Regular	1,705,591,616	NZ
51	Ansell	SKYN Original	1,705,743,116	NZ
	Ansell		, , ,	NZ
52		Lifestyles- Regular	1,701,081,516	
53	Ansell	Lifestyles- Zero	AK017A04	NZ
54	Ansell	Lifestyles-Party Variety- Snake Skin Textured	1,512,032,416	NZ
55	Ansell	Lifestyles – Party Variety - O'Max	1,603,780,316	NZ
56	Ansell	Lifestyles – Party Variety -Tutti Frutti	1,607,192,816	NZ
57	Ansell	Lifestyles – Party Variety -Warm/Cool	1,605,512,216	NZ
58	Ansell	Lifestyles - Party Variety - Glow in the Dark	AGP610A	NZ
59	Ansell	SKYN-Elite	1,702,093,216	NZ
60	Ansell	Lifestyles – Assorted – Banana Bump Studded	1,701,142,016	NZ
61	Ansell	Lifestyles – Assorted –Sonic Strawberry Ribbed	1,606,011,716	NZ
62	Ansell	Lifestyles – Assorted – Berry Blast Smooth	1,601,151,216	NZ
63	Ansell	Lifestyles – Assorted – Vanilla Thriller Smooth		NZ
		·	1,512,101,016	
64	Ansell	Lifestyles – Assorted – Choc Ripple Ribbed	1,610,022,416	NZ
65	Ansell	Lifestyles – Assorted – Mintensity Studded	1,512,050,316	NZ
66	Ansell	Lifestyles – Assorted – Sonic Berry Ribbed	1,702,111,116	NZ
67	Ansell	Lifestyles – Assorted – Banana Bump Studded	1,612,032,416	NZ
68	Ansell	Lifestyles – Party Mix – Choc Ripple Ribbed	1,611,602,316	NZ
69	Ansell	Lifestyles – Party Mix- Dynamint Studded	1,703,090,316	NZ
70	Ansell	Lifestyles - Party Mix - Warm Smooth	1,612,122,916	NZ
71	Ansell	Lifestyles – Party Mix – Tutti Frutti Smooth	1,608,452,116	NZ
72	Ansell	SKYN- Intense Feel	1,612,813,316	NZ
73	Ansell	SKYN-Extra Lubricated	14707PI16	NZ
74	Hero	Natural	7365	NZ
75	Hero	Ultra Thin	7367	NZ
76	GLYDE	Maxi	PN32503	NZ NZ
77	GLYDE	Slimfit	PN31564	NZ
78	Durex	Intense Stimulating	1,000,214,762	NZ
79	Durex	Performa	1,000,170,528	NZ
80	Durex	Together	1,000,174,855	NZ
81	Durex	Love	1,000,033,421	NZ
82	Durex	Close Fit	1,000,065,555	NZ
83	Durex	Real Feel	1,000,243,647	NZ
84	Four Seasons	Naked – Pink Strawberry	X27150902	NZ
85	Four Seasons	Naked Bubble gum	X31150901	NZ
86	Four Seasons	Naked- Chocolate	X26150903	NZ
87	Four Seasons	Naked – Banana Yellow	X33150902	NZ NZ
88	LELO HEX	Original	X35160502	NZ
89	Four Seasons	Extra Strength	X31141101	NZ
90	Four Seasons	Glow	16 N3253	NZ
91	Sir Richard's	Ultra-Thin	13 N3983	NZ
92	GLYDE	Vanilla	PV27711	NZ
93	GLYDE	Strawberry	PS21131	NZ
94	GLYDE	Wild Berry	PW25701	NZ
95	GLYDE	SuperMAX	PN26161	NZ
96	GLYDE	Blueberry	BB23801	NZ
20		· · · · · · · · · · · · · · · · · · ·		
97	GLYDE	Cola	BL28311	NZ

(continued on next page)

(continued)

Sample	Brand	Product	Lot	Purchased
98	GLYDE	Ultra	PN34381	NZ
99	GLYDE	Maxi	PN32503	NZ
100	GLYDE	SlimFit	PN31564	NZ
108	Wet Stuff	Gold Water based personal lubricant	9D729	AUS
109	Ansell	LifeStyles Luxe Silicone-based lubricant	19051507DR	AUS
110	Ansell	Manix Contact	1,606,380,216	CH
111	Ansell	Manix Natural	1,601,012,016	CH
112	Ansell	Manix Orgazmax Plus	1,507,200,316	CH
113 114	Ansell Ansell	Manix Endurance Manix Fraise Gourmande	1,409,151,416 1,408,211,216	CH CH
115	Ansell	Manix Fraise Gournande Manix Xtra Pleasure	1,411,082,316	CH
116	FairSquared	Sensitive dry	0953BA14228	CH
117	FairSquared	Original	0796IF14180	CH
118	FairSquared	Max Perform	1491PF13371	CH
119	Migros	M-Budget	IUR-010	CH
120	Migros	Cosano Regular	OUR-097	CH
121	Migros	Cosano Sensual	OUR-176	CH
122	Migros	Cosano Feeling 0.05 mm	OUR-149	CH
123	Coop	Prix Garantie	1594ZF16451	CH
124	Ceylor	Gold	160831P	CH
125	Ceylor	Thin Sensation	1626B2S	CH
126	Ceylor	Non-Latex UltraThin	660801U	CH
127	Ceylor	Strawberry	153332P	CH
129	Durex	Silver	1,000,176,796	CH
130	Durex	Strawberry	1,000,271,385	CH
131	GLYDE	Silver Lubricated Condom 53 mm	PN34261	AUS
132	GLYDE	Premier Water Based Lubricant	(B) 51,711	AUS
133	Ansell	SKYN Maximum Performance Lubricant	19051103JP	AUS
134	Durex	Play Massage 2 in 1	447X1	AUS
135	Four Seasons	Lush Lubricant with Aloe Vera	9B722	AUS
136	Four Seasons	Nature Lubricant Vegan Friendly	9A716	AUS
137	Durex	KY Jelly	488X2	AUS
138	Four Seasons	Massage Oil	BG668	AUS
139	Durex	Play Feel, Pleasure Gel	213X4	AUS
140	Ansell	LifeStyles Silky Smooth Water Based Lubricant	19,042,006	AUS
141	Durex	Naturals Intimate Gel	919Y3	AUS
142	Astroglide	Naturally Derived Liquide	A011874	AUS
154	Multi-Gyn	Active Gel	1,033,164	AUS
155	Summer's eve	Feminine Wash Sensitive Skin	0217H0161	AUS
156	Femfresh	Feminine deodorant spray	(B)1079351	AUS
157	Femfresh	Daily Intimate Wash	101,702,163	AUS
158	Vagisil	Soothing Oatmeal Cream	Z19D126	AUS
159	Canesten	Intimate Discomfort Cool Cream Gel	GPO1KPL	AUS
160	Vagisil	Intimate Wash Fresh Plus	B91653	AUS
161	Dermeze	Moisturising Cream	B15866	AUS
162	Essentials	Barrier Cream	1,081,235	AUS
163	Rosken	Intensive Moisture Hand Cream	B14519	AUS
164	Love My Ink	Tattoo Cream	30,789	AUS
165	Sasmar	Classic Personal Lubricant	160271C	AUS
166	Sensuous	Smooth & Warming Lubricant	8,433,821	AUS
167	Sensuous	Frenzy Extreme Pleasure Gel for Women	870	AUS
168	Ansell	Skyn Intimate Moments	19031301DR	AUS
169	Ansell Piur Mod	Skyn Intense Feel Non-Latex Condoms	1,905,013,316	AUS
170	Pjur Med	Vegan Glide Intimate Personal Lubricant	32965-01	AUS
171	Astroglide	Diamond Silicone Gel Personal Lubricant	A011819	AUS
172 173	Astroglide Astroglide	Gel Personal Lubricant	A011395 A010316	AUS AUS
173	Astroglide	Personal Lubricant and Massage Oil Waterproof Silicone Liquid	A010316 A011415	AUS
174	Astroglide	Warming Liquid Personal Lubricant	A011415 A011326	AUS
176	Astroglide	Water based Personal lubricant	A011326 A011494	AUS
176	Astroglide	Strawberry Liquid Personal Lubricant	A011494 A011452	AUS
177	Astroglide Four Seasons	Naked Black Condom	Z07180301	AUS
178 179	Four Seasons Four Seasons	Naked Black Condom Stubbed & ribbed Stimulating condoms	18 N0363	AUS
180	Checkmate	ExtraSensitive Lubricated Condoms	1,809,122,816	AUS
181	Durex	Play Perfect Glide	1,809,122,816 541X1	AUS
182	Durex	Climax Stimulating Gel	2231 W7	AUS
	Durex	Comfort XL	1,000,417,937	AUS
183				

Appendix B: Preprocessing methods and effect on the percentage of explained variance

Table B1. preprocessing methods applied on the dataset to enhance the discrimination of the samples. Legend: PC = principal component; SNV = standard normal variation; MSC = Multiple Scattering Correction; Der = derivative.

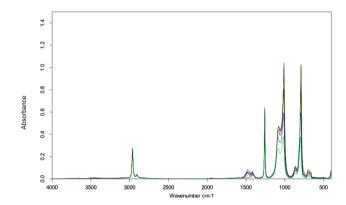
	Percentage of explained variance				
Pre-processing	PC1	PC2	PC3	Total	
Raw data	90	3	3	96	
SNV	84	6	5	95	
MSC	91	8	0	99	
Der1	63	19	5	87	
Der2	46	17	12	75	
SNV + Der1	61	21	6	88	
Der1 + SNV	59	13	8	80	
SNV + Der2	45	19	10	74	
Der2 + SNV	53	13	6	72	
MSC + Der1	95	2	1	98	
Der1 + MSC	82	14	2	98	
MSC + Der2	78	10	3	91	
Der2 + MSC	95	4	0	99	

Appendix C: Investigation of brands, models and lot number discrimination (Figs. C1-C3).

Appendix D: Visual discrimination of infrared spectra and subgroups division (Fig. D1).

Appendix E: Comparison of the discrimination parameters obtained for the five discrimination algorithms tested.

Table E1. Comparison of the performance of the 5 simple-class classification methods, after cross validation on the raw dataset. Aims is to distinguish silicone and non-silicone-based samples. Results are presented for the validation using samples present in the model. Realised on the whole dataset. SVM a linear function, SVM2 uses a second-degree polynomial one and SVM3 uses the radial basis function to compute the data.



 $\begin{tabular}{ll} Fig. \ C1. \ Illustration \ of the \ variation \ within \ different \ models \ of \ condom \ from \ Durex \ brand. \end{tabular}$

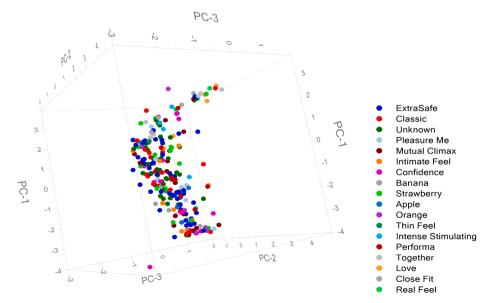


Fig. C2. 3-dimensional score's plot obtained on condom samples from Durex brand, classified according to the different models.

Total samples n 831 831 831 831 True Pos. a 604 605 629 605 False Positive b 25 24 0 24 False Negative c 30 24 7 50 True Negative d 172 178 195 152 Correct Classification Rate (a + d)/n 0,934 0,942 0,992 0,911 Misclassification Rate (b + c)/n 0,066 0,058 0,008 0,089 Sensibility a/(a + c) 0,953 0,962 0,989 0,924 Specificity d/(b + d) 0,873 0,881 1,000 0,864 False Negative Rate c/(a + c) 0,047 0,038 0,011 0,076 False Positive Predictive Power a/(a + b) 0,960 0,962 1,000 0,962 Negative predictive Power d/(c + d) 0,851 0,881 0,965 0,752	SVM3
False Positive b 25 24 0 24 False Negative c 30 24 7 50 True Negative d 172 178 195 152 Correct Classification Rate $(a + d)/n$ 0,934 0,942 0,992 0,911 Misclassification Rate $(b + c)/n$ 0,066 0,058 0,008 0,089 Sensibility a/ $(a + c)$ 0,953 0,962 0,989 0,924 Specificity d/ $(b + d)$ 0,873 0,881 1,000 0,864 False Negative Rate c/ $(a + c)$ 0,047 0,038 0,011 0,076 False Positive Rate b/ $(b + d)$ 0,127 0,119 0,000 0,136 Positive predictive Power a/ $(a + b)$ 0,960 0,962 1,000 0,962	831
False Negative c 30 24 7 50 True Negative d 172 178 195 152 Correct Classification Rate $(a + d)/n$ 0,934 0,942 0,992 0,911 Misclassification Rate $(b + c)/n$ 0,066 0,058 0,008 0,089 Sensibility a/(a + c) 0,953 0,962 0,989 0,924 Specificity d/(b + d) 0,873 0,881 1,000 0,864 False Negative Rate c/(a + c) 0,047 0,038 0,011 0,076 False Positive Rate b/(b + d) 0,127 0,119 0,000 0,136 Positive predictive Power a/(a + b) 0,960 0,962 1,000 0,962	604
True Negative d 172 178 195 152 Correct Classification Rate $(a + d)/n$ 0,934 0,942 0,992 0,911 Misclassification Rate $(b + c)/n$ 0,066 0,058 0,008 0,089 Sensibility $a/(a + c)$ 0,953 0,962 0,989 0,924 Specificity $d/(b + d)$ 0,873 0,881 1,000 0,864 False Negative Rate $c/(a + c)$ 0,047 0,038 0,011 0,076 False Positive Rate $b/(b + d)$ 0,127 0,119 0,000 0,136 Positive predictive Power $a/(a + b)$ 0,960 0,962 1,000 0,962	25
Correct Classification Rate $(a+d)/n$ $0,934$ $0,942$ $0,992$ $0,911$ Misclassification Rate $(b+c)/n$ $0,066$ $0,058$ $0,008$ $0,089$ Sensibility $a/(a+c)$ $0,953$ $0,962$ $0,989$ $0,924$ Specificity $d/(b+d)$ $0,873$ $0,881$ $1,000$ $0,864$ False Negative Rate $c/(a+c)$ $0,047$ $0,038$ $0,011$ $0,076$ False Positive Rate $b/(b+d)$ $0,127$ $0,119$ $0,000$ $0,136$ Positive predictive Power $a/(a+b)$ $0,960$ $0,962$ $1,000$ $0,962$	39
$\begin{array}{llllllllllllllllllllllllllllllllllll$	163
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0,923
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	0,077
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	0,939
False Positive Rate $b/(b+d)$ 0,127 0,119 0,000 0,136 Positive predictive Power $a/(a+b)$ 0,960 0,962 1,000 0,962	0,867
Positive predictive Power $a/(a+b)$ 0,960 0,962 1,000 0,962	0,061
	0,133
Negative modified Power 4/(a + 4) 0.951 0.901 0.005 0.752	0,960
Negative predictive Power $u/(c+u)$ 0,851 0,861 0,965 0,752	0,807
General diagnostic Power $(b+d)/n$ 0,237 0,243 0,235 0,212	0,226
Training accuracy 93,38 94,22 99,16 91,1	92,3
Validation accuracy – – 98,55 91,09	92,05

Table E2. Comparison of the performance of the 5 simple-class classification methods, after cross validation on the raw dataset. Aims is to distinguish condom from other samples Results are presented for the validation using samples present in the model. Realised on the whole dataset. SVM a linear function, SVM2 uses a second-degree polynomial one and SVM3 uses the radial basis function to compute the data.

Measure		LDA	QDA	SVM1	SVM2	SVM3
Total samples	n	831	831	831	831	831
True Pos.	a	633	633	634	633	633
False Positive	b	1	1	0	1	1
False Negative	c	1	1	1	18	1
True Negative	d	196	196	196	179	196
Correct Classification Rate	(a + d)/n	0,998	0,998	0,999	0,977	0,998
Misclassification Rate	(b + c)/n	0,002	0,002	0,001	0,023	0,002
Sensibility	a/(a+c)	0,998	0,998	0,998	0,972	0,998
Specificity	d/(b+d)	0,995	0,995	1,000	0,994	0,995
False Negative Rate	c/(a + c)	0,002	0,002	0,002	0,028	0,002
False Positive Rate	b/(b+d)	0,005	0,005	0,000	0,006	0,005
Positive predictive Power	a/(a + b)	0,998	0,998	1,000	0,998	0,998
Negative predictive Power	d/(c+d)	0,995	0,995	0,995	0,909	0,995
General diagnostic Power	(b + d)/n	0,237	0,237	0,236	0,217	0,237
Training accuracy		99.76	99.76	99,8	97.71	99.76
Validation accuracy		_	_	99.75	97.71	99.75

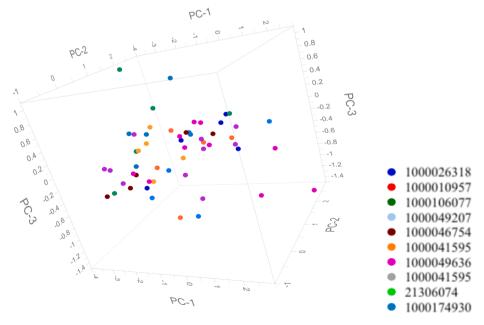


Fig. C3. 3-dimensional score's plot obtained on sample coming from a same brand, same model but presenting different lot number condom samples. Different colors and symbols stands for different lot numbers samples.

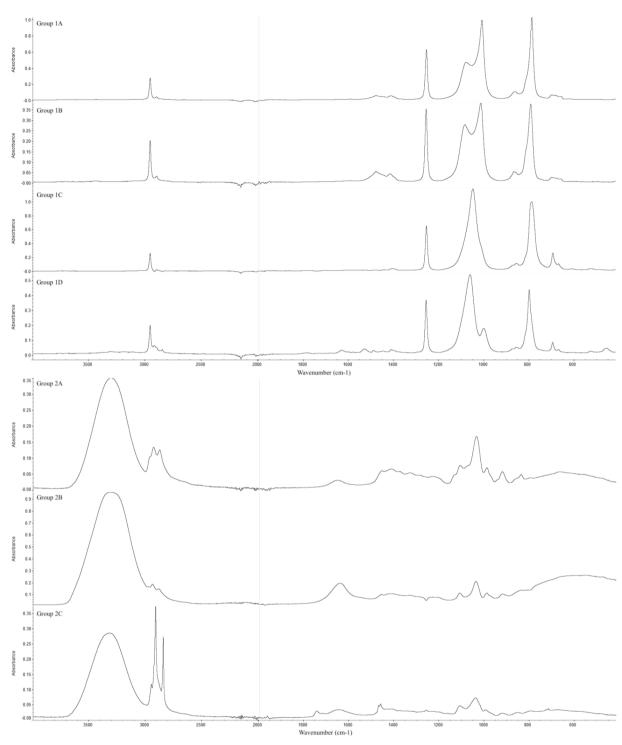


Fig. D1. .

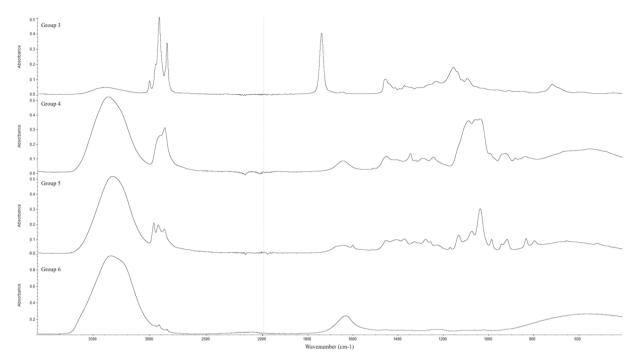


Fig. D1. (continued).

References

- E.N. O'Neal, S.H. Decker, C. Spohn, K. Tellis, Condom use during sexual assault, J. Forensic Leg. Med. 20 (6) (2013) 605–609.
- [2] R.D. Blackledge, M. Vincenti, Identification of polydimethylsiloxane lubricant traces from latex condoms in cases of sexual assault, J. Forensic Sci. Soc. 34 (4) (1994) 245–256.
- [3] T. Coyle, N. Anwar, A novel approach to condom lubricant analysis: In-situ analysis of swabs by FT-Raman Spectroscopy and its effects on DNA analysis, Sci. Justice 49 (1) (2009) 32–40.
- [4] G.P. Campbell, A.L. Gordon, Analysis of Condom Lubricants for Forensic Casework, J. Forensic Sci. 52 (3) (2007) 630–642.
- [5] Z. Shen, J.J. Thomas, G. Siuzdak, R.D. Blackledge, A Case Study on Forensic Polymer Analysis by DIOS-MS: The Suspect Who Gave Us the SLIP®, J. Forensic Sci. 49 (5) (2004) 1–8.
- [6] M. Monzo, The use of the FTIR for the analysis of PDMS in the context of condom traces, Université de Lausanne, 2017.
- [7] P. Brauner, N. Gallili, A Condom—the Critical Link in a Rape, J. Forensic Sci. 38 (5) (1993) 13529J, https://doi.org/10.1520/JFS13529J.
- [8] EWCA Crim 1891, Regina v. Andrew Nicholas Malkinson, 2006.
- [9] P. Maynard, K. Allwell, C. Roux, M. Dawson, D. Royds, A protocol for the forensic analysis of condom and personal lubricants found in sexual assault cases, Forensic Sci. Int. 124 (2-3) (2001) 140–156.
- [10] B. Woffinden, The Nicholas cases: casualties of justice, Bojangles Books, Great Britian, 2016.
- [11] EWCA Crim., Regina v. Andrew Nicholas Malkinson, 2006.
- [12] L.-L. Cho, K.-B. Huang, Identification of condom Lubricants by FT-IR Spectroscopy, Forensic Sci. Int. 11 (2012) 33–40.
- [13] R.D. Blackledge, Viscosity Comparisons of Polydimethylsiloxane Lubricants Latex Condom Brands via Fourier Self-Deconvolution of their FT-IR Spectra, J. Forensic Sci. 40 (1995) 467–469.
- [14] Robert Bradshaw, Rosalind Wolstenholme, Leesa Susanne Ferguson, Chris Sammon, Kerstin Mader, Emmanuelle Claude, Robert D. Blackledge, Malcolm R. Clench, Simona Francese, Spectroscopic imaging based approach for condom identification in condom contaminated fingermarks, Analyst 138 (9) (2013) 2546, https://doi.org/10.1039/c3an00195d.
- [15] C.A. Burnier, W. van Bronswijk, G. Massonnet, Comparison of spectroscopic methods in detection of silicone-based condom lubricants evidence, Anal. Methods. (2020) 10.1039.C9AY02471A. Doi: 10.1039/C9AY02471A.
- [16] Julianne Wolfe, David L. Exline, Characterization of Condom Lubricant Components Using Raman Spectroscopy and Raman Chemical Imaging, J. Forensic Sci. 48 (5) (2003) 2003084.
- [17] National Research Council (U.S.), National Research Council (U.S.), National Research Council (U.S.), eds., Strengthening forensic science in the United States: a path forward, National Academies Press, Washington, D.C, 2009.
- [18] Suzanne Bell, Sunita Sah, Thomas D. Albright, S. James Gates Jr., M. Bonner Denton, Arturo Casadevall, A call for more science in forensic science, Proc. Natl. Acad. Sci. U.S.A. 115 (18) (2018) 4541–4544.

- [19] President's Council of Advisors on Science and Technology, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, Washington, D.C, 2016.
- [20] Brooke Baumgarten, Mark Marić, Lauren Harvey, Candice M. Bridge, Preliminary classification scheme of silicone based lubricants using DART-TOFMS, Forensic Chem. 8 (2018) 28–39.
- [21] Mark Maric, Candice Bridge, Characterizing and classifying water-based lubricants using direct analysis in real time® time of flight mass spectrometry, Forensic Sci. Int. 266 (2016) 73–79.
- [22] Mark Maric, Lauren Harvey, Maren Tomcsak, Angelique Solano, Candice Bridge, Chemical discrimination of lubricant marketing types using direct analysis in real time time-of-flight mass spectrometry: Discrimination of lubricant types using DART-TOFMS, Rapid Commun. Mass Spectrom. 31 (12) (2017) 1014–1022.
- [23] Céline Burnier, Geneviève Massonnet, Pre-analytical considerations of condom traces: A review of composition, background, transfer and persistence, Forensic Sci. Int. 302 (2019) 109861, https://doi.org/10.1016/j.forsciint.2019.06.019.
- [24] Céline Burnier, Geneviève Massonnet, Sally Coulson, David DeTata, Kari Pitts, Characterization and classification of water-based compounds in condoms and personal hygiene products using GC-MS, Forensic Sci. Int. 317 (2020) 110513, https://doi.org/10.1016/j.forsciint.2020.110513.
- [25] D. Lambert, C. Muehlethaler, L. Gueissaz, G. Massonnet, Raman analysis of multilayer automotive paints in forensic science: measurement variability and depth profile: Raman analysis of multilayer automotive paints in forensic science, J. Raman Spectrosc. 45 (2014) 1285–1292, https://doi.org/10.1002/jrs.4490.
- [26] Danny Lambert, Cyril Muehlethaler, Pierre Esseiva, Geneviève Massonnet, Combining spectroscopic data in the forensic analysis of paint: Application of a multiblock technique as chemometric tool, Forensic Sci. Int. 263 (2016) 39–47.
- [27] Cyril Muehlethaler, Geneviève Massonnet, Patrick Buzzini, Influence of the shaking time on the forensic analysis of FTIR and Raman spectra of spray paints, Forensic Sci. Int. 237 (2014) 78–85.
- [28] Cyril Muehlethaler, Genevieve Massonnet, Pierre Esseiva, The application of chemometrics on Infrared and Raman spectra as a tool for the forensic analysis of paints, Forensic Sci. Int. 209 (1-3) (2011) 173–182.
- [29] Cyril Muehlethaler, Geneviève Massonnet, Pierre Esseiva, Discrimination and classification of FTIR spectra of red, blue and green spray paints using a multivariate statistical approach, Forensic Sci. Int. 244 (2014) 170–178.
- [30] Sandra E. Spencer, Sin Young Kim, Seoung Bum Kim, Kevin A. Schug, Matrix-assisted laser desorption/ionization-time of flight-mass spectrometry profiling of trace constituents of condom lubricants in the presence of biological fluids, Forensic Sci. Int. 207 (1-3) (2011) 19–26.
- [31] Yasmine Moustafa, Candice M. Bridge, Distinguishing sexual lubricants from personal hygiene products for sexual assault cases, Forensic Chem. 5 (2017) 58–71.
- [32] Pamela Tozzo, Andrea Gabbin, Caterina Politi, Marta Da Pian, Luciana Caenazzo, Valerio Causin, Combined Statistical Analyses of Forensic Evidence in Sexual Assault: A Case Report and Brief Review of the Literature, J Forensic Sci 65 (5) (2020) 1767–1773.
- [33] Céline Burnier, Geneviève Massonnet, Forensic analysis of condom traces: Chemical considerations and review of the literature, Forensic Sci. Int. 310 (2020) 110255, https://doi.org/10.1016/j.forsciint.2020.110255.